

# POR UMA ABORDAGEM DINÂMICA DOS PROCESSOS FÔNICOS

Adelaide Silva\*  
Vera Pacheco\*\*  
Leonardo Oliveira\*\*

## Introdução

**E**ste texto reúne três estudos – todos apresentados no 4.º CelSul<sup>1</sup> – que têm um mesmo objetivo: argumentar, a partir de dados do português brasileiro (PB), a favor de uma abordagem dinâmica dos processos nesses tratados. Tais dados, somados a outros igualmente analisados no Laboratório de Fonética Acústica e Psicolinguística Experimental (Lafape) do Instituto de Estudos da Linguagem (Unicamp), confirmam a existência de fenômenos fônicos de natureza gradiente, isto é, cuja realização se manifesta através de um contínuo físico. Tais processos, longe de serem meras consequências da implementação mecânica da fala, pertencem ao que Ladefoged (1975) chama

\* Laboratório de Fonética Acústica e Psicolinguística Experimental, Universidade de Campinas; Departamento de Lingüística, Letras Clássicas e Vernáculas, Universidade Federal do Paraná.

\*\* Laboratório de Fonética Acústica e Psicolinguística Experimental, Universidade de Campinas.

1 Realizado na UFPR nos dias 16 e 17 de novembro de 2000.

de “fonética lingüística”, uma vez que fazem parte da gramática da língua, pois são condicionados, por exemplo, pela estrutura prosódica do enunciado no qual ocorrem.

Constituindo a gramática da língua, esses fenômenos devem ser representados nela – pressupõe-se, portanto, ou ao menos da maneira como Fonética e Fonologia são tradicionalmente concebidas,<sup>2</sup> que devam ser representados por algum modelo de análise fonológica. Entretanto, a natureza gradiente de tais fenômenos impede que qualquer modelo fonológico baseado em primitivos teóricos de natureza estática – tal como os traços distintivos – represente os fenômenos de maneira adequada.

Aderimos, então, no Lafape, aos modelos dinâmicos de produção de fala, em especial à Fonologia Articulatória de Browman e Goldstein (1986, 1989, 1990, 1992). Tomando como primitivo o gesto articulatório, uma unidade dotada de tempo intrínseco, esse modelo permite tratar a gradiente verificada em alguns processos fônicos em termos do *timing* e da magnitude gestuais.

Argumentaremos que somente através de um modelo dinâmico conseguiremos dar conta dos fenômenos que vimos observando. Os dados relativos à espirantização da vibrante, investigados por Adelaide Silva, fornecem evidência que suporta uma tal argumentação e colocam em cheque quaisquer tentativas de representação do fenômeno no arcabouço de modelos fonológicos como a Geometria de Traços. Também evidencia a natureza gradiente de processos fônicos a investigação de Vera Pacheco acerca da epêntese do *glide* [j] diante de [s] não palatalizado. Finalmente, o estudo de Leonardo Oliveira aponta para a natureza gradiente da coarticulação perseveratória em seqüências consoante-vocal. Ressalte-se que, neste momento, interessa-nos expor argumentos favoráveis à nossa tese. Propostas de como tratar os fenômenos analisados à luz de um modelo dinâmico como a Fonologia Articulatória (FAR) são matéria para trabalhos posteriores.

## Róticos em início de palavra

Silva (1999) mostra que, em início de palavra no PB, a realização dos róticos pode se dar ao longo de um contínuo físico, isto é, em dialetos onde os

2 Referimo-nos à concepção oriunda do *The sound pattern of English*, segundo a qual à Fonética caberia apenas a implementação mecânica dos sons representados pelo módulo Fonológico, concepção essa que, de certa maneira, foi herdada pelos modelos fonológicos posteriores, inclusive os ditos não-lineares.

falantes ainda produzem a vibrante alveolar, em alternância com a fricativa velar, podem ocorrer “sons intermediários” a esses dois, as chamadas “vibrantes espirantizadas”. A autora nota ainda que a espirantização da vibrante varia em função da força de fronteira adjacente a /r/, tal que quanto mais forte a fronteira, mais o segmento tende a vibrante. Inversamente, quanto mais fraca essa fronteira, mais o segmento tende a uma fricativa. O condicionamento prosódico sobre esse processo atesta sua natureza gramatical e, portanto, torna necessária uma representação fonológica do mesmo.

Partindo desses dados, Adelaide Silva argumenta, neste estudo, que, a despeito de que a espirantização gradual da vibrante deva ser representada fonologicamente, modelos como a Geometria de Traços, na versão de Clements e Hume (1995), não conseguem tratar de fenômenos como esse. Na verdade, sequer conseguem dar conta dos segmentos que constituem a classe das líquidas. Curiosa e paradoxalmente, os autores advogam para seu modelo o poder de realizar um elo entre a representação cognitiva da fala, ou seja, a Fonologia, e sua manifestação física, a Fonética, através dos traços distintivos, por eles definidos como entidades psicológicas dotadas de realização acústica e articulatória (Clements e Hume, 1995, p. 245).

Para Silva, o principal problema dessa versão, no que concerne à impossibilidade de tratamento da gradiente verificada na realização das vibrantes, ou mesmo de um tratamento para as líquidas, parece ser a dificuldade que essa versão da Geometria de Traços tem em lidar com o modo de articulação (D'Angelis, 1998), haja vista que também para o tratamento das nasais a versão de Clements e Hume (1995) se mostra limitada (D'Angelis, 1998). Como decorrência dessa dificuldade, os próprios autores explicitam sua incerteza quanto ao lugar que o traço [lateral] ocuparia na estrutura interna dos segmentos.<sup>3</sup> Além disso, há o fato de que essa versão sequer prevê como seria a estrutura interna dos róticos.

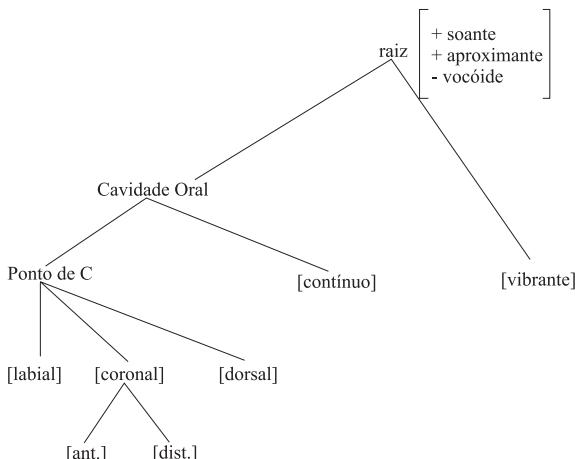
Isso tudo já seria razão para argumentar contra a existência de uma representação possível para os róticos dentro dessa versão da Geometria de Traços. No entanto, ainda podemos tentar propor uma representação que, ao mesmo tempo em que preserva a coerência interna do modelo, “refina-o”. Para

<sup>3</sup> CLEMENTS e HUME (1995) têm duas alternativas: atrelar o traço [lateral] sob o traço [coronal], o que implicaria que todos os segmentos laterais são coronais, o que não é o caso, como eles mesmos apontam, ou atrelar [lateral] sob o nó raiz, opção preferida pelos autores. Para uma argumentação mais detalhada em favor dessa última opção, ver Clements e Hume (1995, p. 293).

tanto, é necessário: 1) diferenciar os róticos das demais classes de segmentos;<sup>4</sup> 2) representar a oposição /r/~/F/; 3) captar a identidade de classe entre segmentos vibrantes e fricativos<sup>5</sup>.

Logo, a proposta de Silva para a representação genérica dos róticos é a seguinte:

(a) *Róticos*



As principais diferenças entre esta representação e aquela que Clements e Hume (1995) propõem para as consoantes são: 1) é descartada a existência de um traço [lateral]; 2) a oposição entre laterais e róticos é dada pelo traço [continuuo]; 3) é previsto um traço [vibrante], que reúne róticos e os diferencia entre si (ao menos aqueles que demandam vibração de um articulador para serem produzidos). Esse traço nem sequer era cogitado em Clements e Hume (1995) e segue a concepção do traço [vibrante] de Ladefoged (1975).

Deve ficar claro que não está em discussão toda a estrutura interna dos segmentos estabelecida por Clements e Hume (1995); tentamos apenas “enxugar” essa representação, recorrendo a um inventário mínimo de traços. Aliás, o único traço inserido na estrutura segmental é [vibrante], o que não significa um

4 Várias línguas há em que róticos se distinguem dos demais fonemas. A esse respeito, ver *The handbook of the International Phonetic Association*. Também nessa referência, é possível encontrar exemplos de línguas que fazem uso da distinção /r/~/F/.

5 Essa identidade é explícita, por exemplo, em processos diacrônicos do PB ou em processos sincrônicos, como o sândi externo do dialeto carioca, que faz com que a fricativa final seja produzida como *tap* diante de vogais.

“inchaço” do modelo ou tampouco a exigência de um processamento cognitivo mais complexo.

O que Silva ressalta sobre a estrutura proposta acima, primeiramente, é que a classe das líquidas é já caracterizada pelos traços que constituem o nó raiz: [+ soante, + aproximante, - vocóide] isola essa classe das obstruintes – [- soante, - aproximante, - vocóide] –, das aproximantes – [- soante, + aproximante, - vocóide] – e dos vocóides [- soante, - aproximante, + vocóide].

Quanto à diferenciação entre laterais e róticos, ela se dá pelo traço [contínuo] e não por [lateral], com base em análises como a de Spencer (1984), as quais suportam a visão de que [lateral] pode ser suprimido da estrutura segmental, sem prejuízo da representação dessa classe, já que o traço [contínuo], mais operacional nos processos fonológicos das línguas, dá conta dessa tarefa. Conseqüentemente, por esta proposta, segmentos especificados positivamente para [contínuo] são laterais;<sup>6</sup> aqueles especificados negativamente para esse traço são róticos.

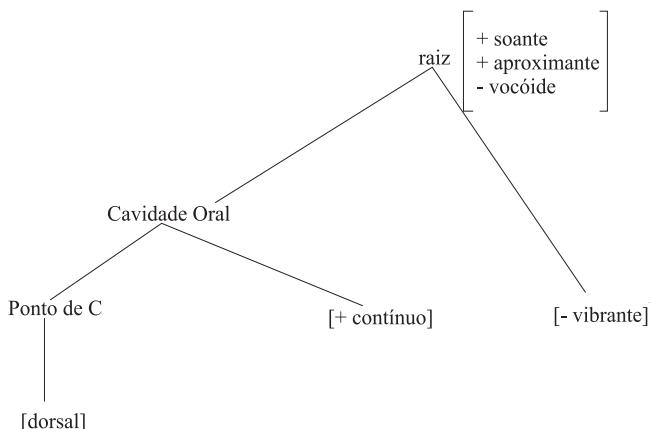
Para representar a oposição /r/~/F/, esta alternativa conta com duas estruturas segmentais possíveis, diferenciadas exclusivamente pelo valor do traço [vibrante]: para a representação da vibrante alveolar, esse traço é especificado positivamente; ao contrário, para a representação do tap, temos [- vibrante].

Na estrutura de uma vibrante uvular, ou de uma vibrante bilabial, por sua vez, o que muda, relativamente à representação (a), é o ponto de articulação. Logo, uma vibrante uvular é também [+ vibrante] e, no entanto, [dorsal], enquanto uma vibrante bilabial, é [labial]. A vantagem de uma tal representação, de acordo com Silva, é que ela capta as relações entre todas as vibrantes – inclusive as que não ocorrem em português.

Resta ainda verificar se a representação atende à exigência de Maia (1982) e capta também a semelhança entre vibrantes e suas variantes fricativas, como a fricativa velar [x], largamente disseminada como variante da vibrante alveolar na maior parte do território brasileiro. Como mencionado acima, a classe das líquidas é caracterizada no nó raiz, através da especificação dos traços que as constituem. Assim, fricativas diferem de vibrantes pelo traço soante, positivo para estas e negativo para aquelas. Nesta proposta, a representação para a fricativa velar é:

6 A oposição entre as laterais, e.g., entre lateral alveolar e lateral palatal, deve ser dada, para Silva, pelo ponto de articulação: a lateral alveolar é [coronal, + anterior] e a palatal, [coronal, - anterior].

(b) *Fricativa velar*



Notamos, portanto, que o traço vibrante não está restrito à classe das líquidas, podendo inclusive ocorrer na estrutura de obstruintes. Pressupor que as fricativas velares sejam especificadas em sua estrutura interna para o traço [vibrante] capta uma semelhança histórica – além de uma identidade fonológica – entre fricativas e vibrantes, semelhança esta atestada por estudos dialetológicos não só para o português brasileiro (Callou et al., 1997) como também para diversos dialetos do espanhol (Obediente, 1998) e que envolve posteriorização e fricativização da vibrante alveolar na direção da fricativa velar.

Apesar de representar a oposição /r/~/F/ e de captar a identidade de classe entre todos os róticos – vibrantes e fricativos –, devemos reconhecer o caráter essencialmente descritivo e arbitrário desta representação. Além disso, outro grande problema da proposta acima é que ela não dá conta de representar as “vibrantes espirantizadas”.

Embora tais segmentos pertençam à gramática do português, pois decorrem da estrutura prosódica das sentenças nas quais os mesmos ocorrem, não há como serem representados por uma estrutura segmental como a apresentada. E a razão dessa impossibilidade está no fato de que, sendo os traços binários, temos apenas duas alternativas: uma vibrante ou uma fricativa, nunca as duas ao mesmo tempo.

7 O traço [vibrante] é especificado negativamente para a fricativa com base na definição que Ladefoged (1975) dá para esse traço: segundo ele, segmentos maximamente vibrados – como as vibrantes – seriam [+ vibrante]; segmentos minimamente vibrados, [- vibrante].

Poderia ser argumentado que uma maneira de resolver esse impasse seria conceber traços escalares também na Geometria de Traços, analogamente ao que propunha Ladefoged (1975).

Sugerindo a incorporação do detalhe fonético à representação fonológica, esse autor defende a possibilidade de prover a representação fonológica de uma especificação fonética em termos da porcentagem dos valores dos traços distintivos, de modo que cada traço seja concebido como uma escala física. Assim, no caso do traço [vibrante], seria teoricamente possível estabelecer uma escala que se estendesse de [0 vibrante] a [100 vibrante], devendo os valores intermediários dessa escala refletir segmentos intermediários a um maximamente vibrado e outro minimamente vibrado.

No entanto, nas afofonias de /r/ em início de palavra não se verificam porcentagens de fricção, isto é, não há um segmento [85 vibrante] ou [50 fricativo]. Verificamos, ao contrário, uma gradiente entre vibrantes e fricativas, podendo ocorrer n vibrantes espirantizadas entre elas, a depender do grau de ruído fricativo sobreposto. Logo, não é difícil concluir que tal fato demandaria n categorias intermediárias a, por exemplo, [0 fricativa] e [1 fricativa]. Assim, voltamos ao impasse decorrente, ao fim e ao cabo, de um problema inerente aos modelos de análise fonológica, em geral, e resultante do fato de que a relação entre Fonética e Fonologia é essencialmente taxonômica, ou seja, alguns fenômenos são classificados como fonéticos, outros como fonológicos, “de tal forma que o mapeamento dos dois níveis torna-se em princípio arbitrário” (Maia, 1982, p. 2).

## A epêntese do glide [j] diante de [s] não palatalizado

Vera Pacheco debruça-se sobre o processo de epêntese do glide [j] diante de [s] em dialetos em que essa fricativa não se palataliza, como o paulista. Para esse estudo, utilizou gravações do telejornal da TV Cultura de São Paulo, considerando as falas de dois repórteres que não palatalizam a fricativa: JP (sexo masculino) e IR (sexo feminino).

Objetivando comparar o ditongo epentético com o ditongo vernáculo, os enunciados desses repórteres, que continham a palavra *mas* monotongada/ ditongada e a palavra *mais*, foram digitalizados a uma taxa de 16 kHz. Foram extraídos espectrogramas de banda larga dessas palavras, num intervalo de tempo de aproximadamente 300ms, com vistas à padronização.

Pacheco mediou os três primeiros formantes do ditongo a partir do espectro de FFT.<sup>8</sup> Dada, porém, a natureza transicional desse ditongo, não foi possível medir os formantes a partir do estado estacionário, tendo sido adotado o procedimento descrito por Albano (1999), o qual considera, para efeito de mensuração, três períodos à esquerda do ruído da fricativa, na transição do elemento epentético para a consoante.

Por se tratar de um *corpus* em que a taxa de elocução do falante não é controlada e, por isso, apresenta diferentes velocidades de fala, foi desenvolvido, para efeito de análise, o índice de duração ID, resultado da razão entre a duração dos segmentos [a<sup>i</sup>, a<sup>j</sup>, a],<sup>9</sup> (Ds), e a duração total da palavra (Dp).

O ID pode ser calculado, então, pela fórmula

$$ID = \frac{Ds}{Dp}$$

onde: Ds, duração dos segmentos [a<sup>i</sup>, a<sup>j</sup>, a], corresponde à medida de duração do intervalo compreendido entre o ponto de transição da nasal para a vogal e o ponto de transição do *glide* (no caso dos ditongos) ou da vogal (no caso dos monotongos) para a fricativa; e Dp, duração total da palavra (Dp), corresponde à medida de duração do intervalo compreendido entre o início e fim da palavra dentro da sentença. O ID relativiza o efeito da taxa de elocução do falante, permitindo a análise da medida de duração do segmento sem essa interferência.

Para as análises estatísticas, Pacheco empregou o teste t de “*Student*”, com o intuito de se certificar se as médias dos parâmetros mensurados (formantes e duração) dos pares *maʃʃs/maʃʃs*, *mas/maʃʃs* apresentam diferença significativa entre si.<sup>10</sup>

Digitalizados os dados e estabelecida a metodologia de análise, Pacheco partiu, então, à análise propriamente dita. Começando pela análise visual, a pesquisadora comparou imagens spectrográficas (figuras 2.1 e 2.2, abaixo), notando que, apesar de auditivamente idênticas, *maʃʃs* aditivo (figura 2.1) difere ligeiramente, no nível acústico, de *maʃʃs* adversativo e auditivamente ditongado (figura 2.2). Isso permite dizer então que as duas palavras não são homônimos perfeitos no nível acústico.

8 Algoritmo usado para se calcular o espectro de Fourier, que mostra os harmônicos da fundamental. As medidas são obtidas manualmente, a partir da identificação visual de picos num espectro de banda estreita derivado através de uma versão digital rápida da Análise de Fourier.

9 Correspondem respectivamente ao ditongo epentético da adversativa ditongada, ao ditongo vernáculo da aditiva e à vogal monotongada da adversativa.

10 A diferença entre as médias foi considerada significativa para  $p < 0,05$ .

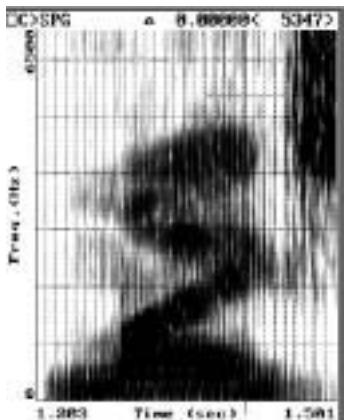


Figura 2.1 - Espectrograma da palavra *mas*, realizada pelo jornalista JP

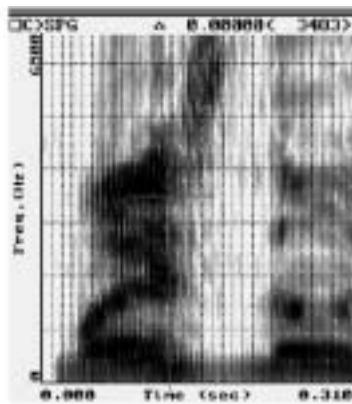


Figura 2.2 - Espectrograma da palavra *mas*, auditivamente ditongo, realizada pelo jornalista JP

A análise visual revela ainda que a “rampa” que caracteriza um ditongo – figura 2.1 – não é tão íngreme no ditongo formado pela epêntese do *glide*, conforme figura 2.2. O aspecto arredondado dessa rampa torna os dois espectrogramas um pouco diferentes entre si. No entanto, essa pequena diferença não impede que o espectrograma da figura 2.2 seja descaracterizado como espectrograma de um ditongo típico, além de não apresentar o formato de um monotongo, também típico. Temos, portanto, uma estrutura intermediária a um ditongo e um monotongo típicos. Não foi possível realizar também, com esse informante, a comparação entre *mas/mafʃs*, em virtude de esse jornalista não realizar a adversativa monotongada. Parece que, para esse falante, o processo de epêntese do *glide* está ocorrendo em todas as situações de sua fala.

Comparando, em seguida, os espectrogramas da aditiva (figura 2.3) com o da adversativa auditivamente ditongada (figura 2.4), obtidos da fala IR (sexo feminino), Pacheco nota a mesma tendência dos movimentos formânticos da adversativa ditongada de JP, fato que a permite afirmar que o processo da epêntese do *glide* [j] ocorre para ambos os sexos.

Diferentemente do jornalista JP, Pacheco detectou, na fala de IR, a ocorrência de adversativas monotongadas, o que lhe permitiu comparar os movimentos formânticos das diferentes realizações da conjunção (ditongada e monotongada) e constatar, efetivamente, a diferença entre as duas realizações.

Pacheco observa também uma diferença substancial dos movimentos formânticos de uma adversativa auditivamente monotongada (figura 2.5) e uma

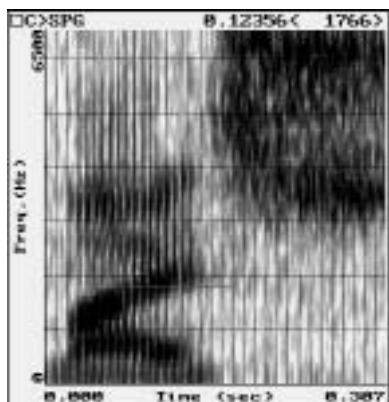


Figura 2.3 - Espectrograma da palavra *mais* falada pela jornalista IR.

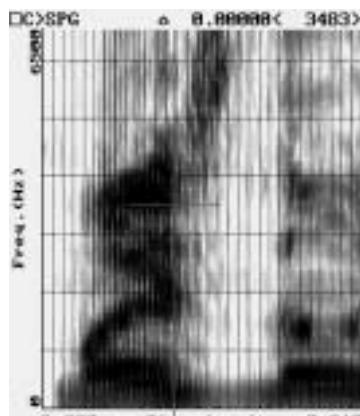


Figura 2.4 - Espectrograma da palavra *mas* realizada auditivamente ditongada pela jornalista IR.

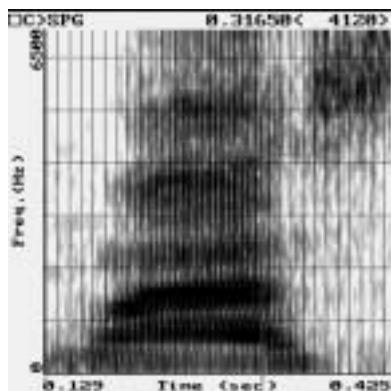


Figura 2.5 - Espectrograma da palavra *mas* realizada auditivamente monotongada pela jornalista IR.

adversativa auditivamente ditongada (figura 2.4, *supra*). Na figura 2.5, não é percebido o movimento de “rampa” que está presente na figura 2.4 (*maʃʃs*), bem como na figura 2.3 (*maʃʃs*). As observações espectrográficas permitem supor que a adversativa, quando ouvida ditongada, possui um movimento que difere sobremaneira da adversativa monotongada e que se aproxima bastante da aditiva. Esse fato corrobora a observação concernente à existência de um segmento intermediário entre ditongo e monotongo típicos, tecida através da comparação entre as figuras 2.1 e 2.2.

Pacheco analisa, em seguida, a freqüência dos três primeiros formantes dos ditongos em *maʃʃs* e *maʃʃs*, calculando as médias desses formantes, que foram extraídos via FFT. Calcula ainda o valor de *p* para as médias da freqüência de cada formante no par *maʃʃs/maʃʃs* nos dados dos dois informantes.

Tabela 2.1- Médias de F1, F2 e F3 de *maʃʃs/maʃʃs* de JP e *mas/maʃʃs* de IR pelo FFT e os valores de *p* (ns) – não significativo – *p* > 0,05; (s) – significativo – *p* < 0,05.

FFT				
Informante	Formante	Ma[ʃʃ]s	Ma[ʃʃ]s	<i>p</i>
JP	F1	291,3	294	0,724 (ns)
	F2	1822,3	1791,4	0,504 (ns)
	F3	2740,3	2492,2	0,001(s)
IR	Formante	Mas	Ma[ʃʃ]s	<i>P</i>
	F1	570	388	0,026 (s)
	F2	1745	2214,7	0,020 (s)
	F3	2395,3	3056	0,0005(s)

A comparação entre os valores de *p* na fala de JP permite que Pacheco observe uma grande proximidade entre F1 e F2, com diferença não significativa para suas médias. Para F3, porém, a diferença entre as médias é significativa (F3, *p* = 0,001), indicando que, para esse informante, o ditongo epentético difere do ditongo estabelecido na língua apenas em F3, fato que corrobora os dados encontrados por Albano (1999), pois a autora encontra também diferença não significativa para dois dos três primeiros formantes (F1 e F3) e diferença significativa para apenas um deles, que no caso de seu estudo é F2.

Os valores de *p* para o par *mas/maʃʃs* realizado pela jornalista IR, por sua vez, apontam para uma diferença significativa para os três formantes, o que permite dizer que, tanto em nível perceptivo quanto em nível acústico, há duas realizações diferentes no par em questão.

Analizando finalmente as medidas de duração, Pacheco nota que elas apresentam resultados distintos para os dois informantes, como mostra a tabela 2.2.

Tabela 2.2 - Índice de Duração (ID) médio das palavras *maʃʃs*, *maʃʃs* e *mas* e respectivos valores de *p* para os informantes IR (sexo feminino) e JP (sexo masculino).

INFORMANTE	ID médio			<i>P</i>
	Ma[ʃʃ]s	Ma[ʃʃ]s	Mas	
IR	0,44	0,57	—	< 0,05 (s)
	0,44	—	0,69	< 0,05 (s)
	0,59	0,57	—	> 0,05 (ns)

A comparação dos valores de ID médio das realizações de *maʃʃs* e de *maʃʃʃs* de IR, com  $p < 0,05$ , indica que os dois segmentos acústicos apresentam diferença significativa de duração. Esses dados somam-se à diferença encontrada nas médias de F3, mostrando que, apesar de auditivamente próximos, esses ditongos apresentam diferenças acústicas.

Pacheco também detectou nos dados de IR diferença significativa ( $p < 0,05$ ) entre os ID médio de *maʃʃs* e *mas*. Os padrões formânticos do ditongo epentético, para essa informante, diferem estatisticamente, ao mesmo tempo, dos padrões formânticos do ditongo estabelecido na língua e dos padrões formânticos da vogal monotongada. Os valores de  $p$  indicam que há ocorrência de três palavras diferentes entre si, acusticamente falando.

No entanto, observando os dados do informante JP, Pacheco ressalva que os resultados são diferentes dos encontrados em IR. Nota, então, que para esse informante não há ocorrência da vogal monotongada. Auditivamente, esse informante não faz distinção entre adversativa e aditiva.

Quando analisamos os valores de ID médio de *maʃʃs* e *maʃʃʃs*, 0,59 e 0,57, respectivamente, praticamente não há diferença entre esses índices, o que é confirmado pelo valor de ( $p > 0,05$ ), que atesta ausência de diferença significativa. Ao contrário da informante IR, o informante JP não faz distinção entre *maʃʃs*, *maʃʃʃs* e *mas*. No que se refere à duração, essas ocorrências são homônimas, acusticamente falando.

Comparando os dados dos dois informantes, pelos dois parâmetros acústicos utilizados na análise – freqüência de formantes e índice de duração médio das palavras –, Pacheco conclui que o processo de ditongação mostra-se em “estágios” diferentes. Enquanto temos, por um lado, IR, que apresenta auditivamente as três realizações com diferença em F3 e em duração, por outro, temos JP, que auditivamente só realiza a vogal ditongada de forma sistemática, apresentando diferença entre as realizações somente para F3.

As diferenças (IR: F3 e duração; JP: F3) e semelhanças (IR: F1 e F2; JP: F1, F2 e duração) acústicas entre *maʃʃs*, *maʃʃʃs* e *mas* encontradas inter e intra falantes são fortes evidências de que o processo de ditongação que vem ocorrendo no PB não é um fenômeno categórico, ou seja, não corresponde à presença versus ausência de ditongação, mas é gradiente. O ditongo epentético apresenta diferenças e semelhanças com o ditongo vernáculo. Esse achado contribui para a reflexão sobre a natureza categórica ou gradiente dos processos fônicos, somando-se aos outros estudos aqui apresentados e também aos estudos atualmente conduzidos no Lafape (Albano et al., 1998).

## Coarticulação consoante-vogal

Entendida, num sentido mais amplo, simplesmente como a influência – decorrente da sobreposição – de um segmento sobre o outro durante a produção da cadeia da fala, a coarticulação pode ser vista tanto como um fenômeno que pertence como um fenômeno que não pertence ao domínio lingüístico. Essa última perspectiva – a de tomar a coarticulação como um fenômeno puramente mecânico – é seguida por modelos tradicionais de análise fonológica, inspirados no clássico SPE.<sup>11</sup>

No entanto, desde o estudo de Öhman (1966), sobre coarticulação de vogal a vogal, sabemos que fenômenos coarticulatórios podem estar sob domínio lingüístico. Nesse sentido, particularmente revelador é o estudo de Whalen (1990), o qual mostra que a coarticulação antecipatória deve estar obrigatoriamente incluída no plano de execução da fala, já que um experimento possibilitou que encontrasse diferenças significativas entre o padrão de coarticulação observado em palavras pronunciadas numa condição normal e o padrão de coarticulação verificado em palavras pronunciadas numa condição em que a locução se iniciava antes de serem conhecidos todos os segmentos da palavra.

Para verificar se também a coarticulação perseveratória é um fenômeno lingüístico – e não meramente mecânico –, ponto que permanece em aberto no estudo de Whalen (1990), Leonardo Oliveira descreve e discute esse tipo de coarticulação entre as vogais /i,a,u/ e as consoantes /p,t,k/, no PB e no inglês americano (IA). Observa ainda, para o PB, a interação da coarticulação perseveratória com a posição acentual – tônica e pós-tônica –, ampliando os resultados encontrados por Aquino (1997).

Para descrever a coarticulação em PB e compará-la diretamente com o IA, foram utilizados três *corpora*:

- 1) logatomas e palavras de estrutura /'laCV/ e /la'CV/,<sup>12</sup> inseridos na frase veículo “Digo *logatoma* baixinho” para o PB. Esses da-

11 Abreviação para *The sound pattern of English* (CHOMSKY, N.; Halle, M. *The sound pattern of English*, New York: Harper & Row, 1968).

12 Nesses *corpora*, a consoante é /p/, /t/ ou /k/ e a vogal, /i/, /a/ ou /u/. Note-se que, a despeito do processo de alofonia de /t/, que gera uma africada [tS] diante de /i/, Oliveira decidiu manter a consoante /t/ nos *corpora* por duas razões: a) esse procedimento possibilita observar três pontos de articulação distintos para verificar o efeito consonantal sobre as vogais; b) os pontos de articulação de [t] e [tS] são próximos, conforme atesta estudo de Stevens e House, 1963 (Cf. STEVENS, K. ; HOUSE, A. S. Perturbation of vowel articulations by consonantal context: acoustic study. *Journal of Speech and Hearing Research*, n. 6, p. 111-128, 1993).

dos foram repetidos dez vezes, em ordem aleatória, por cinco informantes;

- 2) monossílabos tônicos de estrutura /CV/, inseridos na mesma frase-veículo de (1). Os dados foram repetidos cinco vezes, em ordem aleatória, por quatro informantes;
- 3) monossílabos tônicos de estrutura CV, pronunciados isoladamente para o PB e para o IA e repetidos oito vezes, em ordem aleatória, por dois informantes do PB e dois do IA.<sup>13</sup>

Para o PB, os dados foram gravados numa sala acusticamente tratada do Lafape e digitalizados a uma taxa de 16 kHz, usando o recurso de digitalização disponível no CSL (*Computerized Speech Lab*), da Kay Elemetrics, equipamento que também foi utilizado na análise de todos os dados.

Sobre a análise dos dados, cabe observar que foram medidos F1, F2, F3 e F4 de cada vogal, no ponto médio de seu estado estacionário. As medidas foram feitas através do algoritmo LPC (*LinearPredictive Coding*), ajustado a 16 filtros. Em alguns poucos casos, nos quais a análise LPC se mostrava imprecisa, a medição dos formantes foi feita via análise espectral.

Examinando, primeiramente, a vogal /a/, Oliveira observa que os maiores efeitos consonantais sobre a vogal se dão em F2. Há também um efeito discreto, porém inconclusivo, em F<sub>3</sub>. Isso pode ser visualizado, para o corpus de logatomas, na tabela 3.1.

No caso de F2 dos monossílabos e das vogais tônicas dos logatomas, /k/ ocasiona um aumento significativo no valor do formante com relação a /p/, nos dados de todos os informantes. Nas vogais pós-tônicas, por outro lado, /t/ é que causa maior aumento de F2.

Considerando que o correlato articulatório do efeito acústico “aumento de F2” é o avanço do corpo da língua (cf. Lindblom e Sundberg, 1971), esses dados apontam na direção de um provável avanço do corpo da língua para a produção de /a/ seguindo /k/ nos monossílabos e vogais tônicas. Na posição pós-tônica, em que a qualidade da vogal é auditivamente posteriorizada, /t/ é que causa maior aumento de F2.

13 Os dois informantes do IA foram gravados, em condições similares às nossas, no UCLA *Speech Processing and Auditory Perception Laboratory* e os dados desses dois informantes estão disponibilizados na Internet no site <http://www.cscl.ucla.edu/~spapl/>.

A comparação direta entre os dados do PB e do IA evidencia que o valor do F2 de /a/, em ambas as línguas, apresenta diferenças relevantes segundo a consoante precedente, sendo que /k/ sempre eleva o valor do formante em relação a /p/ e varia em relação a /t/.

Tabela 3.1 - Efeito das consoantes precedentes /p, t, k/ no estado estacionário da vogal /a/ tônica nos logatomas. Os maiores efeitos coarticulatórios encontram-se nos quadros sombreados.

Formantes	Informantes				
	Cl	Jl	Gl	Ga	Je
F <sub>1</sub>	$p-v=.046$ $r^2=.203$ $t=p$ $p=k$ $t>k$	$p-v=.924$ $r^2=.005$ $p=t-k$	$p-v=.027$ $r^2=.234$ $p>t$ $t=k$	$p-v=.319$ $r^2=.081$ $p=t-k$	$p-v=.171$ $r^2=.122$ $p=t-k$
F <sub>2</sub>	$p-v=.0001$ $r^2=.826$ $k>t-p$	$p-v=.0001$ $r^2=.595$ $k>t-p$	$p-v=.0001$ $r^2=.536$ $k>t-p$	$p-v=.0007$ $r^2=.417$ $k=t$ $t>p$	$p-v=.0002$ $r^2=.479$ $k=t$ $t>p$
F <sub>3</sub>	$p-v=.0001$ $r^2=.53$ $t>p$ $p=k$	$p-v=.0001$ $r^2=.802$ $t=p$ $p>k$	$p-v=.0001$ $r^2=.524$ $p>k$ $k=t$	$p-v=.0001$ $r^2=.495$ $p=t$ $t>k$	$p-v=.0069$ $r^2=.308$ $p=t$ $t>k$
F <sub>4</sub>	$p-v=.126$ $r^2=.141$ $t=p=k$	$p-v=.605$ $r^2=.036$ $k=p=t$	$p-v=.775$ $r^2=.018$ $p=t=k$	$p-v=.028$ $r^2=.232$ $k=t$ $t=p$ $k>p$	$p-v=.489$ $r^2=.051$ $k=p=t$

Os resultados de F2, portanto, mostram clara semelhança no padrão de coarticulação entre as duas línguas e se mostram coerentes com os resultados já encontrados para as vogais dos logatomas em posição tônica e dos monossílabos isolados.

No caso de F3, por sua vez, os resultados obtidos se mostram inconclusivos, já que apenas um dos informantes apresenta diferenças significativas em relação à consoante precedente (vide tabela 3.1). Além disso, dados de um dos informantes sugerem que a coarticulação perseveratória pode variar no mesmo indivíduo.

Examinando, em seguida, a vogal /u/, Oliveira nota que sobre ela acontece o efeito coarticulatório mais robusto, tal que o valor de F2 é sistematicamente elevado em contexto de /t/, tanto em posição tônica como em posição pós-tônica, para monossílabos isolados e em contexto – aqueles inseridos na frase-veículo “Digo monossílabo baixinho” (cf. (1) do item 1 *supra*) – em ambas as línguas, isto é, PB e IA. A única exceção, como se vê pela tabela 3.2, foi o informante Ga, cujas vogais pós-tônicas apresentaram grande variação no valor de F2.

Para os monossílabos isolados, Oliveira encontrou diferenças entre as duas línguas relativamente a F3: nos informantes do IA não há diferenças significativas em relação à consoante precedente. Para os informantes do PB, a consoante /p/ aparece sempre elevando o valor do F3 de /u/ quando comparada com /k/. Este resultado foi encontrado também para os logatomas, mas não para os monossílabos em contexto. Nesse caso, há uma diferença entre as duas línguas, mas a inconsistência dos dados do PB enfraquece esse resultado, já que possivelmente o padrão não se mantém nem mesmo no PB, dadas as diferenças encontradas para o mesmo indivíduo, variando apenas o *corpus*.

Tabela 3.2 - Efeito das consoantes precedentes /p, t, k/ no estado estacionário da vogal /u/ pós-tônica nos logatomas. Os quadros sombreados indicam os maiores efeitos coarticulatórios.

		Informantes				
Formantes		Cl	J1	Gl	Ga	Je
F <sub>1</sub>	$p-v=.02$	$p-v=.332$	$p-v=.056$	$p-v=.75$	$p-v=.001$	
	$r^2=.249$	$r^2=.078$	$r^2=.206$	$r^2=.133$	$r^2=.371$	
	$p=k$	$p=t=k$	$p=k$	$p=k=t$	$p>t$	
	$k>t$		$k=t$	$p>t$	$t=k$	
F <sub>2</sub>	$p-v=.0001$	$p-v=.0001$	$p-v=.0001$	$p-v=.14$	$p-v=.0001$	
	$r^2=.838$	$r^2=.833$	$r^2=.623$	$r^2=.624$	$r^2=.685$	
	$t>k$	$t>p>k$	$t>k>p$	$t=p=k$	$t>k$	
	$k=p$				$k=p$	
F <sub>3</sub>	$p-v=.0001$	$p-v=.018$	$p-v=.0001$	$p-v=.703$	$p-v=.001$	
	$r^2=.691$	$r^2=.255$	$r^2=.713$	$r^2=.161$	$r^2=.381$	
	$t=p$	$p=t$	$p>t$	$p=k=t$	$p=t$	
	$p>k$	$t=k$	$t=k$		$t>k$	
F <sub>4</sub>	$p-v=.173$	$p-v=.208$	$p-v=.283$	$p-v=.379$	$p-v=.568$	
	$r^2=.121$	$r^2=.109$	$r^2=.124$	$r^2=.383$	$r^2=.040$	
	$t=p=k$	$k=p=t$	$k=p=t$	$t=k=p$	$k=t=p$	

Quanto ao efeito coarticulatório das consoantes sobre a vogal /i/, cabe mencionar que se mostra inconsistente, pois a coarticulação varia tanto de acordo com o informante como de acordo com a consoante precedente. Uma possível causa para essa variação pode estar associada a uma tentativa de impedir a proximidade entre F2 e F3 de /i/, pois um aumento sistemático do F2 – tal como observado para as outras duas vogais – poderia afetar a distância entre F2 e F3, fazendo com que ambos fossem percebidos como um único formante (cf. Stevens, 1989), modificando, assim, o reconhecimento auditivo de /i/.

Interessantemente, o efeito coarticulatório encontrado por Oliveira para o PB verifica-se também em outras línguas: Schouten e Pols (1979), observam que, no alemão, F2 de /a/ e /u/ aumenta no contexto de /t/ em relação a /p/.

Também os resultados de Stevens e House (1963), para o IA, são semelhantes aos do PB – devemos notar, porém que, neste caso, o efeito coarticulatório é maior sobre /u/. Essas comparações apontam para uma universalidade da coarticulação perseveratória consoante-vocal, mas dadas as diferenças de metodologia entre os estudos e as diferenças entre o estudo de Oliveira e o de Stevens e House (1963), não podemos afirmar que o fenômeno é completamente mecânico ou que tem algo de lingüístico.

Qualquer que seja a natureza da coarticulação, contudo, uma hipótese dinâmica modelaria mais adequadamente o fenômeno. Isso porque modelos dessa natureza poderiam, em princípio, explicar: 1) a gradiência; 2) a possível não-linearidade da coarticulação; 3) o aspecto lingüístico da coarticulação, se for o caso. Isso tudo sem a necessidade de se postular dois níveis de descrição – o fonético e o fonológico –, o que torna a hipótese dinâmica mais parcimoniosa a princípio. Modelos fonológicos tradicionais,<sup>14</sup> como a Geometria de Traços, não prevêem a gradiência encontrada nos resultados. Mesmo com tendências claras em certas direções, a coarticulação perseveratória se apresenta como intrinsecamente variável e, assim, só um modelo fonético/fonológico que leve em conta essa gradiência nos parece favorável para explicar os dados.

## Sumário e discussão

Os dados relativos à espirantização da vibrante de início de palavra e os dados relativos à epêntese do *glide* [j] diante de [s] não palatalizado evidenciam um problema crucial dos modelos fonológicos tradicionais ao tentar lidar com esse tipo de processos fônicos: baseando-se em primitivos de análise categóricos, quaisquer modelos desses prevêem apenas a transformação de uma categoria fônica X noutra categoria fônica Y. Não prevêem, em hipótese alguma, que entre essas duas categorias possa existir um contínuo físico ao longo do qual emergem vários segmentos intermediários. Resulta então que qualquer tentativa de representar processos fônicos gradientes, nesses arcabouços teóricos, torna-se completamente arbitrária, como tentamos mostrar na sessão 1.

Como mencionamos na introdução, é preciso representar esses processos fônicos na gramática da língua, dado que se mostram condicionados por algum fato lingüístico – como a estrutura prosódica do enunciado no qual se inserem (Cf., Albano et al., 1998). Nesse sentido, os únicos modelos capazes de

14 Entendemos por “modelos fonológicos tradicionais” aqueles que se calcam em primitivos teóricos de natureza categórica, como os traços distintivos.

prever a existência de segmentos intermediários a duas categorias fônicas são os chamados “modelos dinâmicos”, como a Fonologia Articulatória de Browman e Goldstein (1986, 1989, 1990, 1992). Tomando como unidade o “gesto articulatório”, dotado de tempo intrínseco, tal modelo assume que a variação de magnitude e de *timing* dos gestos é responsável pela emergência de inúmeros segmentos intermediários a duas categorias. Ora, se há segmentos que tendem a vibrantes ou a fricativas conforme a força da fronteira adjacente a /r/, por um lado, e se há segmentos intermediários a ditongos e monotongos típicos, por outro, a única maneira de representá-los é prever que os gestos articulatórios envolvidos na produção de róticos e [j]s variam de tal modo em magnitude e *timing*, que consequentemente surgem os segmentos intermediários a vibrantes e fricativas, por um lado, e entre ditongo e monotongo, por outro.

Os dados relativos à coarticulação perseveratória reforçam ainda mais nossa filiação aos modelos dinâmicos, pois trazem à baila uma outra vantagem desses modelos: como mostramos, os dados sobre a coarticulação perseveratória sugerem a universalidade desse fenômeno, pois coincidem, ao menos em parte, com os de outras línguas. Embora ainda não tenha ficado claro se essa coarticulação é ou não lingüística, os modelos dinâmicos podem modelá-la suficientemente bem: além de preverem que a gradiente dos processos fônicos emerge automaticamente a partir dos gestos articulatórios, porque estes são dotados de tempo intrínseco, os modelos dinâmicos, como a FAR, consideram que Fonética e Fonologia constituem um único módulo de processamento, diferentemente de outros modelos fonológicos. Essa previsão simplifica sobremaneira a análise, pois já não é mais necessário prever dois níveis de análise e todo o conjunto de regras que transformam o *output* de um nível no *input* do seguinte. Como consequência, temos um modelo mais parcimonioso, o que é um forte argumento para preferi-lo aos demais.

## RESUMO

Dados do português brasileiro fornecem forte evidência acerca da natureza gradiente de alguns processos fônicos. Tais dados sugerem ainda terem natureza lingüística, o que requer sua representação em algum modelo lingüístico. No entanto, os modelos fonológicos – como concebidos tradicionalmente – não conseguem lidar com processos gradientes, dada a natureza categórica de seus primitivos de análise. Conseqüen-

temente, a única maneira de tratar tais processos fônicos é analisá-los à luz dos “modelos dinâmicos”, como a “Fonologia Articulatória”. Neste texto apresentamos evidências do caráter gradiente de alguns processos fônicos, além de descrevê-los e discutir os limites de um modelo como a Geometria de Traços ao tentar lidar especialmente com a espirantização da vibrante.

*Palavras-chave: Fonética acústica, gradiente, modelos dinâmicos.*

## ABSTRACT

Brazilian Portuguese data provide strong evidence towards the gradient nature of some speech phenomena. Moreover, those data strongly suggest that they are linguistic in nature. Consequently, they must be represented in some linguistic framework. On the other hand, phonological models, as traditionally conceived, cannot deal with gradient phenomena, mainly because of the categorical nature of their primitives. So, the only way to deal with such phenomena is analyzing them in light of the so-called “dynamic models”, such as “Articulatory Phonology”. In this text we aim at presenting evidence on the gradient nature of some speech phenomena, as well as describing them and discussing the limits of Feature Geometry in dealing with trills spirantization in particular.

*Key-words: Acoustic phonetics, gradience, dynamic models.*

## REFERÊNCIAS

ALBANO, E. A gestural solution for some glide epenthesis problems. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 14., 1999, San Francisco. *Proceedings...* San Francisco: [s. n.], 1999, p. 1785-1789.

ALBANO, E. et al. Interface Fonética-Fonologia e a interação Prosódia-segmentos. In: SEMINÁRIO GEL, 45., 1998, Campinas. *Anais XLV.*

AQUINO, P. A. *O papel das vogais reduzidas pós-tônicas na construção de um sistema de síntese concatenativa para o português do Brasil.* Campinas, 1997. Dissertação (Mestrado) - IEL/ Unicamp. No prelo.

- BROWMAN, C.; GOLDSTEIN, L. Towards an articulatory phonology. *Phonology Yearbook*, n. 3, p. 219-252, 1986.
- \_\_\_\_\_. Articulatory gestures as phonological units. *Phonology*, n. 6, p. 201-251, 1989.
- \_\_\_\_\_. Tiers in articulatory phonology, with some implicarions for casual speech. In: KINGSTON, J.; BECKMAN, M. E. (Eds.). *Papers in laboratory phonology I: between the grammar and the physics of speech*. Cambridge: The Cambridge University Press, 1990.
- \_\_\_\_\_. Articulatory phonology: an overview. *Phonetica*, n. 49, p. 155-180, 1992.
- CALLOU, D.; MORAES, J. A.; LEITE, Y. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: KOCH, I. V. (Org.). *Gramática do português falado*. Campinas: Editora da UNICAMP, 1997. v. 6.
- CHOMSKY, N.; HALLE, M. *The sound pattern of English*. New York: Harper & Row, 1968.
- CLEMENTS, N.; HUME, E. Internal organization of speech sounds. In: GODSMITH (Ed.). *The handbook of phonological theory*. Cambridge: Blackwell, 1995. p. 245-306.
- D'ANGELIS, W. R. *Traços de modo e modos de traçar geometrias: línguas Macro-Jê & teoria fonológica*. Campinas, 1998. Tese (Doutorado) - IEL/Unicamp, No prelo.
- HERNANDORENA, C. L. M. Introdução à teoria fonológica. In: BISOL, L. (Org.). *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1996. p. 09-86.
- LADEFOGED, P. *A course in phonetics*. New York: Harcourt Brace Jovanovich, 1975.
- LINDBLOM, B.; SUNDBERG, J. Acoustical consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America*, n. 35, p. 1773-1781, 1971.
- MAIA, E. A. M. Evidência portuguesa para uma nova versão do traço rotacionado. In: ENCONTRO NACIONAL DE LINGÜÍSTICA, 1982. *Anais...* [S. l.: s. n.], 1982. No prelo.
- ÖHMAN, S. Coarticulation in VCV utterances: spectrographic measurements. *Journal of the Acoustical Society of America*, n. 39, p. 151-168, 1966.
- OBEDIENTE, E. *Fonética y fonología*. 3. ed. Mérida: Facultad de Humanidades y Educación/ Consejo de Publicaciones de la Universidad de los Andes, 1998.
- SCHOUTEN, M. E. H.; POLS, L. C. W. Vowel segments: aspectral study of coarticulation – part I. *Journal of Phonetics*, n. 7, p. 1-23, 1979.

SILVA, A. Pistas para o condicionamento prosódico sobre a variabilidade de produção de /r/. *Estudos Lingüísticos*, n. 28, p. 682-688, 1999a.

SPENCER. Eliminating the feature [lateral]. *Journal of Linguistics*, n. 20, p. 23-43, 1984.

STEVENS, K. On the quantal nature of speech. *Journal of Phonetics*, n. 17, p. 3-46, 1989.

STEVENS, K.; HOUSE, A. S. Perturbation of vowel articulations by consonantal context: an acoustic study. *Journal of Speech and Hearing Research*, n. 6, p. 111-128, 1963.

WHALEN, D. H. Coarticulation is largely planned. *Journal of Phonetics*, n. 18, p. 3-35, 1990.

